SEQUENCE LISTING

<110>	ENEA-Ente per le Nuove Tecnologie e l'Ambiente Consiglio Nazionale delle Ricerche											
<120>	Method for the preparation of transgenic plants characterised by Geminivirus lasting resistance											
<130>	PCT25622											
<140> <141>	RM2003A000242 2003-05-19											
<150> <151>	RM2003A000242 2003-05-19											
<160>	12											
<170>	PatentIn version 3.2											
<210> <211> <212> <213>	1 630 DNA Geminivirus TYLCSV											
<400> atgcca	1 agat caggtcgttt tagtatcaag gctaaaaatt atttccttac atatcccaaa	60										
tgtgat	ttaa caaaagaaaa tgcactttcc caaataacaa acctacaaac acccacaaac	120										
aaatta	ttca tcaaaatttg cagagaacta catgaaaatg gggaacctca tctccatatt	180										
ctcato	caat tcgaaggaaa atacaattgt accaatcaac gattcttcga cctggtatcc	240										
ccaaco	aggt cagcacattt ccatccgaac attcagggag ctaaatcgag ctccgacgtc	300										
aagtco	tata tcgacaagga cggagatgtt cttgaatggg gtactttcca gatcgacgga	360										
cgatct	gcta ggggaggaca acagacagcc aacgacgctt acgcaaaggc aattaacgca	420										
ggaagt	taagt cgcaggctct tgatgtaatt aaagaattag cgcctagaga ttacgttcta	480										
catttt	tcata atataaatag taatttagat aaggttttcc aggtgcctcc ggcaccttat	540										
gtttct	tcctt ttttatcttc ttctttcgat caagttcctg atgaacttga acactgggtt	600										
tccgag	gaacg tcatggatgc cgctgcgcgg	630										
<210> <211> <212> <213>	630 DNA											
<220> <223>	TYLCSV Rep-210 modified sequence											
<220> <221> <222>	CDS											
<400> atg comet Post 1	2 ct aga tcc gga agg ttt agc atc aaa gct aag aat tac ttc ttg ro Arg Ser Gly Arg Phe Ser Ile Lys Ala Lys Asn Tyr Phe Leu 5 10 15	48										
aca t	ac ccc aag tgt gac tta act aag gag aat gca ttg tcc cag ata • Pagina 1	96										

						,							•			
Thr	Tyr	Pro	Lys 20	Cys	Asp	Leu	Thr	Lys 25	Glu	Asn	Ala	Leu	ser 30	G In	IJе	
act Thr	aac Asn	ttg Leu 35	caa Gln	act Thr	ccc Pro	act Thr	aac Asn 40	aag Lys	ttg Leu	ttc Phe	att Ile	aag Lys 45	att Ile	tgt Cys	agg Arg	144
gaa Glu	ctt Leu 50	cat His	gag Glu	aat Asn	gga Gly	gaa Glu 55	cca Pro	cat His	ctt Leu	cat His	atc Ile 60	ttg Leu	ata Ile	cag Gln	ttc Phe	192
gaa Glu 65	ggc Gly	aag Lys	tat Tyr	aac Asn	tgc Cys 70	acc Thr	aac Asn	caa Gln	cgt Arg	ttc Phe 75	ttt Phe	gac Asp	ctt Leu	gtg Val	tcc ser 80	240
cct Pro	acc Thr	aga Arg	tca Ser	gcc Ala 85	cat His	ttt Phe	cat His	cca Pro	aac Asn 90	atc Ile	cag Gln	ggt Gly	gct Ala	aag Lys 95	tcg Ser	288
agt Ser	tca Ser	gac Asp	gtg Val 100	aag Lys	tca Ser	tac Tyr	att Ile	gac Asp 105	aaa Lys	gac Asp	ggc Gly	gat Asp	gtg Val 110	Ļeu	gag Glu	336
tgg Trp	gga Gly	act Thr 115	Phe	cag Gln	ata Ile	gac Asp	ggt Gly 120	Arg	tcg Ser	gct Ala	aga Arg	gga Gly 125	GIY	cag Gln	caa Gln	384
aca Thr	gct Ala 130	Asn	gat Asp	gca Ala	tac Tyr	gct Ala 135	aag Lys	gct Ala	atc Ile	aac Asn	gct Ala 140	. Сту	tcc Ser	aag Lys	tca Ser	432
cag Gln 145	Ala	ctt Leu	gac Asp	gta Val	atc Ile 150	Lys	gag Glu	tta Leu	gct Ala	cct Pro 155	Arg	gat Asp	tat Tyr	gtt Val	ctt Leu 160	480
cat His	tto Phe	cat His	aac Asr	ato 11e 165	Asn	agc Ser	aat Asn	ttg Leu	gac Asp 170	Lys	gtç Val	tto Phe	caa Gln	gtg Va 175	cca Pro	. 528
ccg Pro	gct	cct Pro	tac Tyr 180	' Val	tca Ser	cct Pro	ttc Phe	tta Leu 185	Ser	tct Ser	tca Ser	tti Phe	gat Asp 190	יוט י	ggtt val	576
cca Pro	gat Asp	gag Gli 19!	ı Lei	gag u Glu	cat His	tgg Trp	gtg Val 200	Sei	gaa Glu	aac Asr	gti Va	t 'ato 1 Mei 20!	r wat	gce Ala	gca a Ala	624
gcg Ala	cgi Ard 210	3			•											630
<2.	LO> L1> L2> L3>	3 210 PRT Art	ific	ial												
<22 <22	20> 23>	Syn	thet [.]	ic C	onsti	ruct										
<40	>00	3														
Me ¹	t Pr	o Ar	g Se	r Gly 5	y Ar	g Pho	e se	r Il	e Ly: 10	s Ala	a Ly	s As	n Ty	r Ph 15	e Leu	
Thi	т ту	r Pr	o Ly 20	s Cy:	s As	p Lei	ı Th	r Ly 25	s Glu Pa	u Ası gina	n Al 2	a Le	u Se 30	r Gl	n Ile	

Thr Asn Leu Gln Thr Pro Thr Asn Lys Leu Phe Ile Lys Ile Cys Arg 35 40 45

Glu Leu His Glu Asn Gly Glu Pro His Leu His Ile Leu Ile Gln Phe 50 60

Glu Gly Lys Tyr Asn Cys Thr Asn Gln Arg Phe Phe Asp Leu Val Ser 65 70 75

Pro Thr Arg Ser Ala His Phe His Pro Asn Ile Gln Gly Ala Lys Ser 85 90 95

Ser Ser Asp Val Lys Ser Tyr Ile Asp Lys Asp Gly Asp Val Leu Glu 100 105

Trp Gly Thr Phe Gln Ile Asp Gly Arg Ser Ala Arg Gly Gln Gln 115 120 125

Thr Ala Asn Asp Ala Tyr Ala Lys Ala Ile Asn Ala Gly Ser Lys Ser 130 135 140

Gln Ala Leu Asp Val Ile Lys Glu Leu Ala Pro Arg Asp Tyr Val Leu 145 150 155 160

His Phe His Asn Ile Asn Ser Asn Leu Asp Lys Val Phe Gln Val Pro 165 170 175

Pro Ala Pro Tyr Val Ser Pro Phe Leu Ser Ser Ser Phe Asp Gln Val 180 185 190

Pro Asp Glu Leu Glu His Trp Val Ser Glu Asn Val Met Asp Ala Ala 195 200 205

Ala Arg 210

<210> 4 <211> 630

<211> 630 <212> DNA

<213> Artificial

<220> <223> TYLCSV Rep-210 modified sequence

<220> <221> CDS

<222> (1)..(630)

48

aca Thr	tac Tyr	ccc Pro	aag Lys 20	tgt Cys	gac Asp	tta Leu	act Thr	aag Lys 25	gag Glu	aat Asn	gca Ala	ttg Leu	tcc ser 30	cag Gln	ata Ile	96
act Thr	aac Asn	ttg Leu 35	caa Gln	act Thr	ccc Pro	act Thr	aac Asn 40	aag Lys	ttg Leu	ttc Phe	att Ile	aag Lys 45	att Ile	tgt Cys	agg Arg	144
gaa Glu	ctt Leu 50	cac His	gag Glu	aat Asn	gga Gly	gaa Glu 55	cca Pro	cat His	ctt Leu	cat His	atc Ile 60	ttg Leu	ata Ile	cag Gln	ttc Phe	192
gaa Glu 65	ggc Gly	aag Lys	tat Tyr	aac Asn	tgc Cys 70	acc Thr	aac Asn	caa Gln	cgt Arg	ttc Phe 75	ttt Phe	gac Asp	ctt Leu	gtg Val	tcc Ser 80	240
cct Pro	acc Thr	aga Arg	tca Ser	gcc Ala 85	.cat His	ttt Phe	cat His	cca Pro	aac Asn 90	atc Ile	cag Gln	ggt Gly	gct Ala	aag Lys 95	tcg Ser	288
agt Ser	tca Ser	gac Asp	gtg Val 100	aag Lys	tca Ser	tac Tyr	att Ile	gac Asp 105	aaa Lys	gac Asp	ggg Gly	gat Asp	gtg Val 110	ctc Leu	gag Glu	336
tgg Trp	gga Gly	act Thr 115	ttt Phe	cag Gln	ata Ile	gac Asp	ggt Gly 120	cga Arg	tcg Ser	gct Ala	aga Arg	gga Gly 125	ggt Gly	cag Gln	caa Gln	384
aca Thr	gca Ala 130	Asn	gat Asp	gca Ala	tac Tyr	gct Ala 135	aag Lys	gct Ala	atc Ile	aac Asn	gct Ala 140	gga Gly	tcc Ser	aag Lys	tca Ser	432
cag Gln 145	Ala	ctt Leu	gac Asp	gta Val	atc Ile 150	aaa Lys	gag Glu	tta Leu	gct Ala	cct Pro 155	agg Arg	gat Asp	tat Tyr	gtt Val	ctt Leu 160	480
cat His	ttc Phe	cat His	aac Asn	atc Ile 165	Asn	agc Ser	aat Asn	ttg Leu	gac Asp 170	aaa Lys	gtg Val	ttc Phe	caa Gln	gtg Val 175	cca Pro	528
ccg Pro	gct Ala	cct Pro	tac Tyr 180	Val	tca Ser	cct Pro	ttc Phe	tta Leu 185	Ser	tct Ser	tca Ser	ttt Phe	gat Asp 190	Gin	gtt Val	576
cca Pro	gat Asp	gag Glu 195	Leu	gag Glu	cat His	tgg Trp	gtg Val 200	ser	gaa Glu	aac Asn	gtt Val	atg Met 205	ASP	gcc Ala	gca Ala	624
	cgt Arg 210															630
<21 <21 <21 <21	.1> .2>	5 210 PRT Arti	fici	al												
<22 <22		Synt	:heti	c Co	nstr	uct										
<40	0>	5					•									•
Met 1	Pro	Arg	, Ser	· Gly 5	/ Arg	Phe	ser	Ile	Lys 10	аТа	Lys	Asn	Tyr	Phe 15	Leu	

Thr Tyr Pro Lys Cys Asp Leu Thr Lys Glu Asn Ala Leu Ser Gln Ile Pagina 4

30

20

25

Thr Asn Leu Gln Thr Pro Thr Asn Lys Leu Phe Ile Lys Ile Cys Arg 35 40 45

Glu Leu His Glu Asn Gly Glu Pro His Leu His Ile Leu Ile Gln Phe 50 60

Glu Gly Lys Tyr Asn Cys Thr Asn Gln Arg Phe Phe Asp Leu Val Ser 65 70 75 80

Pro Thr Arg Ser Ala His Phe His Pro Asn Ile Gln Gly Ala Lys Ser 85 90 95

Ser Ser Asp Val Lys Ser Tyr Ile Asp Lys Asp Gly Asp Val Leu Glu 100 105 a 110

Trp Gly Thr Phe Gln Ile Asp Gly Arg Ser Ala Arg Gly Gly Gln Gln 115 120 125

Thr Ala Asn Asp Ala Tyr Ala Lys Ala Ile Asn Ala Gly Ser Lys Ser 130 135 140

Gln Ala Leu Asp Val Ile Lys Glu Leu Ala Pro Arg Asp Tyr Val Leu 145 150 155 160

His Phe His Asn Ile Asn Ser Asn Leu Asp Lys Val Phe Gln Val Pro 165 170 175

Pro Ala Pro Tyr Val Ser Pro Phe Leu Ser Ser Phe Asp Gln Val 180 185 190

Pro Asp Glu Leu Glu His Trp Val Ser Glu Asn Val Met Asp Ala Ala 195 200 205

Ala Arg 210

<210> 6

<211> 774 <212> DNA

<213> Artificial

<220> <223> TYLCSV Coat Protein modified sequence

<220> <221> CDS <222> (1)..(774)

 $^{<\!400>}$ 6 atg cca aag aga act ggt gat att cta atc tca act ccc gtg tct aag Met Pro Lys Arg Thr Gly Asp Ile Leu Ile Ser Thr Pro Val Ser Lys 1 Pagina 5

48

gtg Val	cgt Arg	agg Arg	aga Arg 20	ctt Leu	aac Asn	ttt Phe	gac Asp	tct Ser 25	ccg Pro	tac Tyr	acc Thr	tct Ser	cgt Arg 30	gca Ala	gct Ala	96
gct Ala	ccc Pro	aca Thr 35	gtc Val	cag Gln	ggc Gly	att Ile	aag Lys 40	agg Arg	cga Arg	tct Ser	tgg Trp	aca Thr 45	Tyr	aga Arg	cct Pro	144
atg Met	tac Tyr 50	agg Arg	aaa Lys	ccg Pro	agg Arg	atg Met 55	tat Tyr	agg Arg	atg Met	tat Tyr	cgt Arg 60	agc Ser	cca Pro	gat Asp	gtg Val	192
cct Pro 65	cct Pro	ggt Gly	tgc Cys	gaa Glu	gga Gly 70	ccc Pro	tgc Cys	aag Lys	gtg Val	caa Gln 75	tcg Ser	tat Tyr	gag Glu	caa Gln	cgt Arg 80	240
gac Asp	gat Asp	gtg Va1	aag Lys	cac His 85	acc. Thr	gga Gly	gtt Val	gtt Val	cgt Arg 90	tgc Cys	gtt val	tct Ser	gat Asp	gtg Val 95	act Thr	288
aga Arg	ggt Gly	tca Ser	ggt Gly 100	Ile	act Thr	cac His	agg Arg	gtg Val 105	gga Gly	aag Lys	cgt Arg	ttc Phe	tgt Cys 110	att Ile	aag Lys	336
tct Ser	att Ile	tac Tyr 115	ata Ile	ttg Leu	ggt Gly	aag Lys	atc Ile 120	tgg Trp	atg Met	gac Asp	gag Glu	aat Asn 125	atc Ile	aag Lys	aaa Lys	384
cag Gln	aat Asn 130	cac His	act Thr	aat Asn	cag Ģln	gtt Val 135	atg Met	ttc Phe	ttt Phe	ctt Leu	gtg Val 140	cga Arg	gat Asp	cga Arg	aga Arg	432
cca Pro 145	Tyr	gga Gly	acc Thr	agc Ser	cca Pro 150	atg Met	gac Asp	ttc Phe	ggc Gly	cag Gln 155	gtg Val	ttt Phe	aat Asn	atg Met	ttc Phe 160	480
gat Asp	aac Asn	gag Glu	cca Pro	tct Ser 165	Thr	gca Ala	act Thr	gtg Val	aaa Lys 170	Asn	gat Asp	ttg Leu	cgt Arg	gat Asp 175	aga Arg	528
tat Tyr	cag Gln	gtg Val	atg Met 180	Arg	aag Lys	ttc Phe	cat His	gca Ala 185	Thr	gtg Val	gtt Val	ggt Gly	ggt Gly 190	י ארס	tct Ser	576
gga Gly	atg Met	aaa Lys 195	Glu	caa Gln	tgt Cys	ctt Leu	ctg Leu 200	Lys	aga Arg	ttc Phe	ttt Phe	aag Lys 205	Ile	aac Asn	act Thr	624
cat His	gtc Val 210	Val	tat Tyr	aac Asn	cac His	cag Gln 215	Glu	caa Gln	gcg Ala	aaa Lys	tat Tyr 220	GIL	ı aat ı Asr	cac His	act Thr	672
gaa Glu 225	Asn	gct Ala	ttg Leu	ttg Leu	tta Leu 230	Tyr	atg Met	gcc Ala	tgt Cys	acc Thr 235	His	gca	tct Ser	aat Asr	cca Pro 240	720
gtt Val	tac Tyr	gca Ala	acg Thr	ctt Leu 245	Lys	atc Ile	cgt Arg	ato	tat Tyr 250	. Phe	tat Tyr	gad Asp	gct Ala	gto Val 255	aca Thr	768
aac Asn	tag	l														774

<210> 7 <211> 257 <212> PRT <213> Artificial

<220> <223> Synthetic Construct

<400> 7

Met Pro Lys Arg Thr Gly Asp Ile Leu Ile Ser Thr Pro Val Ser Lys
1 10 15

Val Arg Arg Leu Asn Phe Asp Ser Pro Tyr Thr Ser Arg Ala Ala 20 25 30

Ala Pro Thr Val Gln Gly Ile Lys Arg Arg Ser Trp Thr Tyr Arg Pro 35 40 45

Met Tyr Arg Lys Pro Arg Met Tyr Arg Met Tyr Arg Ser Pro Asp Val 50 60

Pro Pro Gly Cys Glu Gly Pro Cys Lys Val Gln Ser Tyr Glu Gln Arg 65 70 75 80

Asp Asp Val Lys His Thr Gly Val Val Arg Cys Val Ser Asp Val Thr 85 90 95

Arg Gly Ser Gly Ile Thr His Arg Val Gly Lys Arg Phe Cys Ile Lys
100 105 110

Ser Ile Tyr Ile Leu Gly Lys Ile Trp Met Asp Glu Asn Ile Lys Lys 115 120 125

Gln Asn His Thr Asn Gln Val Met Phe Phe Leu Val Arg Asp Arg Arg 130 135 140

Pro Tyr Gly Thr Ser Pro Met Asp Phe Gly Gln Val Phe Asn Met Phe 145 150 155 160

Asp Asn Glu Pro Ser Thr Ala Thr Val Lys Asn Asp Leu Arg Asp Arg 165 170 175

Tyr Gln Val Met Arg Lys Phe His Ala Thr Val Val Gly Gly Pro Ser 180 185 190

Gly Met Lys Glu Gln Cys Leu Leu Lys Arg Phe Phe Lys Ile Asn Thr 195 200 205

His Val Val Tyr Asn His Gln Glu Gln Ala Lys Tyr Glu Asn His Thr 210 215 220

Glu Asn Ala Leu Leu Leu Tyr Met Ala Cys Thr His Ala Ser Asn Pro 225 230 235 240 Val Tyr Ala Thr Leu Lys Ile Arg Ile Tyr Phe Tyr Asp Ala Val Thr 245. 250 255

Asn

<210: <211: <212: <213:	> ' > 1	3 447 DNA Artif	icia	1		•										
<220: <223:		TYLCSV Rep 130 sequence														
<220 <221 <222	>	CDS (51).	. (44	3)											Š .	
<222	> · >	misc_ (231) Point	(2	31) .	n fr	om C	(Re	p-21	.0 wi	1d-t	ype)	to	Т			•
<220 <221 <222 <223	> >	misc_ (233) Point	(2	(33	n fr	om C	(Re	p 21	.O wi	ld-t	ype)	to	G			
<400 ggat		8 cct <u>c</u>	gata	cttt	g ag	tgtc	cccc	gat	tcag	jaac	gaca	.gcaa	aa a M	let P	ca Pro	56
aga Arg	tca Ser	ggt Gly 5	cgt Arg	ttt Phe	agt Ser	atc Ile	aag Lys 10	gct Ala	aaa Lys	aat Asn	tat Tyr	ttc Phe 15	ctt Leu	aca Thr	tat Tyr	104
ccc Pro	aaa Lys 20	tgt Cys	gat Asp	tta Leu	aca Thr	aaa Lys 25	gaa Glu	aat Asn	gca Ala	ctt Leu	tcc ser 30	caa G1n	ata Ile	aca Thr	aac Asn	152
cta Leu 35	caa G1r	aca Thr	ccc Pro	aca Thr	aac Asn 40	aaa Lys	tta Leu	ttc Phe	atc Ile	aaa Lys 45	att Ile	tgc Cys	aga Arg	gaa Glu	cta Leu 50	200
cat His	gaa Glu	aat Asn	ggg Gly	gaa Glu 55	cct Pro	cat His	ctc Leu	cat His	att Ile 60	ttg Leu	atc Ile	caa Gln	ttc Phe	gaa Glu 65	gga Gly	248
aaa Lys	tac Tyi	aat Asn	tgt Cys 70	acc Thr	aat Asn	caa Gln	cga Arg	ttc Phe 75	ttc Phe	gac Asp	ctg Leu	gta Val	tcc ser 80	cca Pro	acc Thr	296
agg Arg	tca Sei	gca Ala 85	cat His	ttc Phe	cat His	ccg Pro	aac Asn 90	att Ile	cag Gln	gga Gly	gct Ala	aaa Lys 95	tcg Ser	agc Ser	tcc Ser	344
gac Asp	gtg Va 100	aag Lys)	tcc Ser	tat Tyr	atc Ile	gac Asp 105	aag Lys	gac Asp	gga Gly	gat Asp	gtt Val 110	ctt Leu	gaa Glu	tgg Trp	ggt Gly	392
act Thr 115	tto Pho	cag Gln	atc Ile	gac Asp	gga Gly 120	cga Arg	tct Ser	gct Ala	agg Arg	gga Gly 125	gga Gly	caa Gln	cag Gln	aca Thr	gcc Ala 130	440

447

tga attc

<210> 130 <211>

<212> PRT

<213> Artificial

<220>

<223> Synthetic Construct

<400> 9

Met Pro Arg Ser Gly Arg Phe Ser Ile Lys Ala Lys Asn Tyr Phe Leu 1 10 15

Thr Tyr Pro Lys Cys Asp Leu Thr Lys Glu Asn Ala Leu Ser Gln Ile 20 25 30

Thr Asn Leu Gln Thr Pro Thr Asn Lys Leu Phe Ile Lys Ile Cys Arg 35 40 45

Glu Leu His Glu Asn Gly Glu Pro His Leu His Ile Leu Ile Gln Phe 50 60

Glu Gly Lys Tyr Asn Cys Thr Asn Gln Arg Phe Phe Asp Leu Val Ser 65 70 75 80

Pro Thr Arg Ser Ala His Phe His Pro Asn Ile Gln Gly Ala Lys Ser 85 90 95

Ser Ser Asp Val Lys Ser Tyr Ile Asp Lys Asp Gly Asp Val Leu Glu 100 105 110

Trp Gly Thr Phe Gln Ile Asp Gly Arg Ser Ala Arg Gly Gln Gln 115 120 125

Thr Ala 130

<210> 10 30

<211> <212> DNA

Artificial

<220> <223>

Primer

<220>

misc_feature (1)..(30) <221>

Primer for PCR C4 mutagenesis

<400> 10

ctcatctcca tattttgatc caattcgaag

<210> 11

30

<211> <212> <213>	30 DNA Arti	ficial					
<220> <223>	Prim	er					
<220> <221> <222> <223>	(1).	_feature .(30) ner for PCR	C4 mutagene	esis			
<400> cttcga	11 attg	gatcaaaata	tggagatgag				30
<210> <211> <212> <213>	12 774 DNA Gemi	inivirus TYL	.csv	•		e La Sala	
<400> atgccg	12 aagc	gaaccggcga			tctcgaaggt	- tcgtcgaaga	60
					ccactgtcca	•	120
cgtcga	tcat	ggacttacag	gcccatgtat	cgaaagccgc	ggatgtacag	aatgtacaga	180
agccct	gatg	tacctccggg	ttgtgaaggt	ccctgtaaag	tgcagtcgta	cgagcagcgt	240
gatgac	gtca	agcataccgg	tgttgtgcgt	tgtgttagtg	atgtaactag	gggttctggt	300
attact	cata	gagttggtaa	acgtttttgt	atcaagtcaa	tttatatatt	aggaaagatt	360
tggatg	gatg	aaaacataaa	aaaacaaaat	catactaacc	aagtgatgtt	tttccttgtt	420
cgagac	cgaa	ggccttatgg	aactagtcct	atggattttg	gtcaagtttt	taacatgttt	480
gataat	gaac	ccagtactgc	tacggtgaag	aacgacttac	gggataggta	tcaagtaatg	540
aggaag	jtttc	atgctacggt	tgttggaggt	ccgtcaggga	tgaaggagca	gtgtttgctg	600
aagaga	itttt	ttaaaattaa	tacccatgta	gtttataatc	accaagagca	ggcgaagtat	660
gaaaat	cata	ctgagaatgc	cttgttattg	tatatggctt	gtactcatgc	ttctaaccca	720
atatea	-acta	cottoaaaat	acutatttat	ttttatgatg	ctataacaaa	ttaa	774